Moral Hazard and Bargaining Power*

Dominique Demougin and Carsten Helm†

Forthcoming: German Economic Review

Abstract

We introduce bargaining power in a moral hazard framework where parties are risk-neutral and the agent is financially constrained. We show that the same contract emerges if the concept of bargaining power is analyzed in either of the following three frameworks; in a standard P-A framework by varying the agent's outside opportunity, in an alternating offer game, and in a generalized Nash bargaining game. However, for sufficiently low levels of the agent's bargaining power, increasing it marginally does affect the equilibrium in the Nash bargaining game, but not in the P-A model and in the alternating offer game.

JEL Classification: D2; D8; L14

Keywords: Principal-agent model; bargaining power; moral hazard.

---

*We wish to thank Oliver Fabel, Eberhard Feess, Roland Strausz, Anja Schöttner, Veikko Thiele and two anonymous referees for valuable comments.

†Demougin: Humboldt University Berlin, School of Business and Economics, Spandauer Str. 1, D-10178 Berlin, Germany, demougin@wiwi.hu-berlin.de. Helm: Technical University Darmstadt, Department of Law and Economics, Marktplatz 15, D-64283 Darmstadt, Germany, Tel: +49 6151 16-2095, Fax: +49 6151 16-6043, helm@vwl.tu-darmstadt.de.

1 Introduction

A standard assumption in the Principal-Agent model is that the principal makes 'take it or leave it' offers to the agent. However, for most real world problems both parties hold some bargaining power. Furthermore, Pitchford (1998) recently pointed out that for a large class of cases where the agent has limited liability the distribution of bargaining power between principal and agent affects the joint surplus generated by the contract. This may have important policy implications. For example, the design of labor market institutions plays an essential role in determining the distribution of bargaining power between employers and employees.

To analyze the implications of bargaining power, some authors have compared only the two extreme situations where either the Principal or the Agent can make 'take it or leave it' offers (e.g. Mookherjee and Ray 2002). By contrast, Pitchford (1998) considers also intermediate allocations of bargaining power, which he represents by varying the agent's reservation utility in a standard P-A model. A different approach is adopted by Balkenborg (2001), who uses the Nash bargaining solution to analyze a similar moral hazard problem. Alternatively, one could analyze bargaining power in an alternating offer game with model hazard. None of the cited papers discusses the relationship between these different options, and the reasons for adopting their particular approach. The present paper aims to fill this gap.

For the case of risk-neutral parties and a financially constrained agent we show that the same set of contracts arises from varying the agent's reservation utility in a P-A model, the discount factors in an alternating offer game à la Rubinstein (1982), or the bargaining power coefficient in a Nash bargaining game. Since our moral hazard model gives rise to a concave Pareto frontier, this equivalence does not come as a surprise. However, due to moral hazard and the liability limit the relationship between the different ways to represent bargaining power – through the reservation utility, discount factors and bargaining power coefficients – is not one-to-one as in a 'standard' bargaining game. In particular, variations in the reservation utility or in discount factors may have no effect on bargaining outcomes, while changes in the bargaining power coefficient always do so.

In the following, we analyze the three approaches to model bargaining power.

2 The P-A model with varying outside options

We consider a P-A environment with risk-neutral parties. The value of output for the principal is $v(e)$, where $e \in \mathbb{R}^+$ is the agent's effort associated with costs $c(e)$. We impose standard requirements, assuming that $v(e)$ is increasing and concave with $v'(0) = \infty$, while $c(e)$ is increasing and convex.

None of the above variables is verifiable, resulting in moral hazard. The principal and the agent observe a contractible binary signal $s \in \{0, 1\}$, where $s = 1$ is a favorable signal (see Milgrom 1981).[1] Denoting $p(e) \equiv \Pr\{s = 1|e\}$, we assume $p'(e) > 0, p''(e) < 0$.[2] Due to the informational assumptions, contracts will also be binary. We denote with F the fixed payment and with b the bonus. Finally, we assume that the agent is financially constrained. Specifically, we require $F, F + b \geq 0$ so that the first-best is not always obtainable.

In the standard P-A model, the principal's expected utility is

$$
\begin{aligned}
\pi(\bar{u}) \equiv \max_{\{F,b,e\}} \quad & v(e) - [F + bp(e)] \text{ subject to} && (1)\\
& bp'(e) = c'(e) && (2)\\
& F \geq 0 && (3)\\
& F + bp(e) - c(e) \geq \bar{u}, && (4)
\end{aligned}
$$

where (2) is the incentive compatibility condition, (3) the liability limit given $b > 0$, and (4) the constraint on the agent's utility. Substituting b from (2) yields the expected bonus $B(e) = c'(e)p(e)/p'(e)$, which we assume to be convex. This assumption ensures that the first-order condition of the Lagrangian is sufficient.

Upon varying the agent's reservation utility, we get the following result.[3]

Proposition 1 *In the P-A model, the principal's utility is decreasing concave in the agent's reservation utility $\bar{u}$. For low values of $\bar{u}$, the optimal contract has $F = 0$, the agent extracts rent and effort is constant at the second best e^{**}. For high values of $\bar{u}$, the optimal contract implements first-best effort e^* and has $F = \bar{u} + c(e^*) - B(e^*)$. For intermediate values of $\bar{u}$, $F = 0$, and effort is increasing in $\bar{u}$.*

Proof. From the Lagrangian of the principal's optimization problem we get the first-order conditions w.r.t. e and F,

$$
\begin{aligned}
v'(e) - B'(e) + \mu(B'(e) - c'(e)) &= 0 && (5)\\
-1 + \lambda + \mu &= 0, && (6)
\end{aligned}
$$

where λ and μ are the Lagrangian multipliers for the limited liability and participation constraint, respectively.

[1] This is a generalized version of the problem in Pitchford (1998). Note that the assumption $s \in \{0, 1\}$ is without loss of generality, as in the risk-neutral agency problem all relevant information from a mechanism design point of view can be summarized by a binary statistic (see, e.g., Kim 1997).

[2] These conditions guarantee that the agent's problem is well behaved. They are equivalent to considering binary signals satisfying MLRC and CDFC.

[3] For similar results see Pitchford (1998) or Demougin and Fluet (2001), and for an adverse selection context Inderst (2002).

There are three cases. With $\lambda = 1, \mu = 0$, by complementary slackness $F = 0$ and the standard second-best effort, e^{**}, obtains from (5). This can only arise if $B(e^{**}) \geq c(e^{**}) + \bar{u}$. If the inequality is strict, the agent extracts a rent and small variations in $\bar{u}$ leave e^{**}, π^{**} and u^{**} unaffected.

As $\bar{u}$ increases, the constraint on the agent's utility must become binding at some point and $\lambda, \mu > 0$. Thus, fixed payments remain at 0 and effort follows from the binding constraint on the agent's utility. Implicitly differentiating w.r.t. $\bar{u}$ yields $e_{\bar{u}} = [B'(e) - c'(e)]^{-1} > 0$, where the sign follows from the definition of $B(e)$ and the curvature assumptions. From the envelope theorem we have $\pi_{\bar{u}} = -\mu < 0$. Totally differentiating (5) and rearranging yields $\mu_{\bar{u}} > 0$ so that for intermediate values of $\bar{u}$ the Pareto frontier is decreasing and concave. With $\mu = 1, \lambda = 0$ effort attains the social optimum by (5) and F follows from the binding participation constraint.[4] □

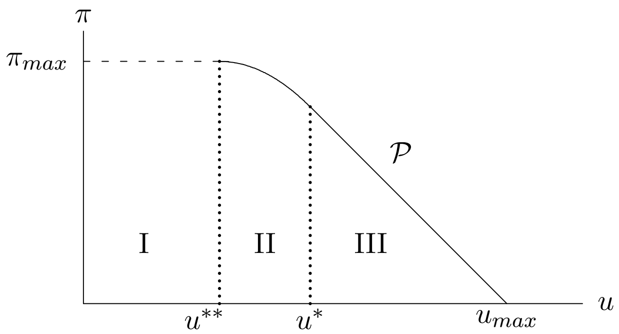

Figure 1: The constrained Pareto frontier

Figure 1 depicts the constrained Pareto frontier, denoted $\mathcal{P}$, of the set of possible utility pairs (π, u). In particular, let $c \equiv \{F, b, e\}$ be an incentive compatible contract that satisfies the agent's financial constraint. Then $(\pi, u) \in \mathcal{P}$ if and only if it is implementable by a contract c, and there exists no other contract c that is Pareto preferred. For future references, we denote the set of contracts c that lead to utility pairs on the constrained Pareto frontier by $\mathcal{C} \equiv \{c : u(c), \pi(c) \in \mathcal{P}\}$. Observe that there exists a one-to-one mapping between $\mathcal{C}$ and $\mathcal{P}$.

When $\bar{u} < u^{**}$, the principal offers a contract yielding utility u^{**} for the agent, who extracts a rent $u^{**} - \bar{u}$. Accordingly, the dashed line above region I does not belong to $\mathcal{P}$ since it cannot arise from a contract c. Raising $\bar{u}$ reduces the agent's rent until it falls to zero. When $\bar{u} \geq u^{**}$, increases in the agent's reservation utility must be compensated by either raising b or F. Increasing b is initially advantageous as it raises effort (region II). Once

[4] This third case can never arise in the framework analyzed by Pitchford due to the binary nature of $v(e)$ in his model.

effort is first best, a further increase in $\bar{u}$ is best compensated by lump sum transfers (region III).

3 Bargaining game with alternating offers

We now consider the case where the principal and the agent bargain over incentive compatible contracts that satisfy the agent's financial constraint $F, F+b \geq 0$. We model the bargaining process as an alternating offer game. In round 1, the principal offers a contract. If the agent accepts, the contract is implemented. If he declines, the game proceeds to a second round where the agent proposes a contract. If the principal accepts, the contract is implemented. If she rejects, the game continues in the same manner with alternating offers. We assume that the principal and the agent are impatient and denote with $\delta_p, \delta_a \in [0,1]$ their respective discount factors.

For the moment, assume that there exists a unique subgame perfect equilibrium. Perfection requires that the parties offer only contracts $c \in \mathcal{C}$. Accordingly, we can describe contract offers by points on the constrained Pareto frontier $\mathcal{P}$. Furthermore, from the previous section we know that $\mathcal{P}$ is a concave function $\pi(u)$ whose domain is the interval $[u^{**}, u_{max}]$ and range the interval $[0, \pi_{max}]$. Note that, unlike the standard bargaining game, the agent's impasse point, 0, does not belong to the domain of $\pi(u)$.

To determine the equilibrium, suppose the game were to attain period 3 in which the principal offers $u_p \in [u^{**}, u_{max}]$ to the agent. Going one period back, the agent will match the principal's present value of her period 3 utility, i.e. offer $\pi_a = \delta_p \pi(u_p)$. Again going back one period, the principal will offer $u_p = \max\{u^{**}, \delta_a u(\pi_a)\}$ as we know from proposition 1 that it can never be optimal to offer less than u^{**} (see Figure 1). Accordingly, in contrast to the standard alternating offer game there are situations where the principal offers the agent more than his reservation utility $\delta_a u(\pi_a)$. By stationarity of the game, we find upon substitution

$$u_p = \max\{u^{**}, \delta_a u(\delta_p \pi(u_p))\}. \tag{7}$$

To prove existence of a unique subgame perfect equilibrium and to characterize the fixed point, we map this expression in Figure 2. In period 3 the agent's utility will always be between u^{**} and u_{max}. Since $u_{max} \geq \delta_a u(\delta_p \pi(u_{max}))$, the point D lies below the diagonal. Furthermore, the slope of the function $\delta_a u(\delta_p \pi(u_p))$ is

$$0 \leq \frac{\partial \delta_a u(\delta_p \pi(u_p))}{\partial u_p} = \delta_a u'(\cdot) \delta_p \pi'(u_p) \leq 1. \tag{8}$$

To verify the inequality, observe that from $u_p \leq u(\delta_p \pi(u_p))$ and the concavity of $\pi(u)$, we obtain $u'\pi' \leq 1$. Moreover, u' and π' are both negative and $\delta_p, \delta_a \leq 1$. Hence there are two cases. If $\delta_a u(\delta_p \pi(u^{**}))$ lies above the diagonal (like point A), then we have an interior fixed point like u_p. Otherwise, $\delta_a u(\delta_p \pi(u))$ is like BD and the fixed point is u^{**}. Variations in the discount factors shift the curves. However, this affects only interior fixed points. Unlike the standard alternating offer game, small variations in the discount factors have no effect if the equilibrium offer is u^{**}.

Finally, it is straightforward to show that any point along the constrained Pareto frontier can be represented by different profiles of the discount factors. For example, suppose $\delta_p = 0$. With $\delta_a = 0$ the agent receives u^{**}, while with $\delta_a = 1$, he receives $u(0) = u_{\max}$. Thus by continuity as δ_a rises from 0 to 1, the agent's utility must take all the values between u^{**} and u_{max}.

We obtain the following result, where the second claim follows from the fact that $\delta_a u(\delta_p \pi(u_p))$ in (7) increases in δ_a and decreases in δ_p.

Proposition 2 *The alternating offer game has a unique subgame perfect equilibrium in contracts. Moreover, let $(\delta_p^{**}, \delta_a^{**})$ be a profile of discount factors such that equilibrium contract offers lead to $(\pi(u^{**}), u^{**})$. Then for all $\{(\delta_p, \delta_a) : \delta_p \geq \delta_p^{**}, \delta_a \leq \delta_a^{**}\}$ equilibrium utility remains at $(\pi(u^{**}), u^{**})$.*

4 The Nash bargaining solution

Binmore, Rubinstein, and Wolinsky (1986) have shown that the standard bargaining process with alternating offers can be approximated by the Nash bargaining solution. We extend their result to the current moral hazard set up. To do so, we initially hold bargaining power α constant and maximize the Nash bargaining product

$$[F + bp(e) - c(e)]^{\alpha}[v(e) - F - bp(e)]^{1-\alpha}, \tag{9}$$

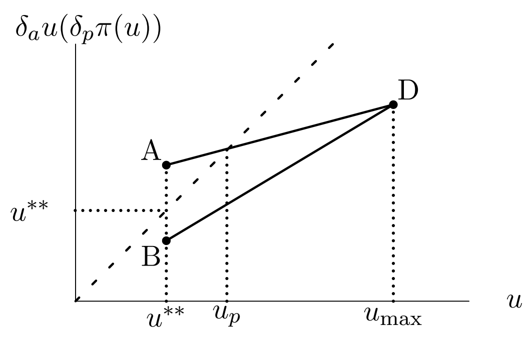

Figure 2: Fix points in alternating offer game

with respect to contracts that are ex-post incentive compatible and satisfy the agent's financial constraint.[5] With $\alpha = 0$, the Nash bargaining product equals the principal's utility, and with $\alpha = 1$ the agent's utility. Obviously, for these extreme cases the solution corresponds to the boundary points of the constrained Pareto frontier (see Figure 1). For $\alpha \in (0, 1)$, the corresponding Lagrangian becomes

$$\mathcal{L}(e, F, \xi) = \alpha \ln\left[F + B(e) - c(e)\right] + (1 - \alpha) \ln\left[v(e) - F - B(e)\right] + \xi F,$$

with first-order conditions

$$\frac{\alpha\,(B'(e) - c'(e))}{F + B(e) - c(e)} + \frac{(1-\alpha)(v'(e) - B'(e))}{v(e) - F - B(e)} = 0 \tag{10}$$

$$\frac{\alpha}{F + B(e) - c(e)} - \frac{(1-\alpha)}{v(e) - F - B(e)} + \xi = 0. \tag{11}$$

When $\xi = 0$, the first best solution obtains since from substituting (11) into (10)

$$\frac{\alpha(B'(e) - c'(e))}{F + B(e) - c(e)} + \frac{\alpha(v'(e) - B'(e))}{F + B(e) - c(e)} = 0, \tag{12}$$

which implies $v'(e) = c'(e)$. Furthermore, (11) can be solved for

$$\alpha = \frac{F}{v(e^*) - c(e^*)} + \frac{B(e^*) - c(e^*)}{v(e^*) - c(e^*)}, \tag{13}$$

where the second term on the r.h.s defines a critical level of bargaining power α_c. For $\alpha \geq \alpha_c$ the first best effort obtains and any increase in bargaining power results in a larger F. Moreover, from (13) as α approaches 1 so that the agent has the entire bargaining power, he extracts all the profit and attains utility u_{max}.

When $\xi > 0$, complementary slackness implies $F = 0$. Implicitly differentiating (10) then yields $de/d\alpha > 0$. Note that this inequality is strict so that variations in α always affect the equilibrium contract. Finally, as α approaches 0 the standard second best obtains.

In conclusion, as the agent's bargaining power α goes from 0 to α_c, effort increases from the second best to the first best level and F remains at 0. As α increases further, effort stays at the first best level and F adjusts. Finally, observe that even though it is desirable from a welfare point of view to attain first best effort, the principal will not willingly relinquish bargaining power as it lowers her utility.

Proposition 3 *The constrained Pareto frontier $\mathcal{P}$ (or, equivalently, the set of contracts $\mathcal{C}$) can be defined alternatively as the solution of (i) the P-A*

[5]Note that participation is guaranteed by construction of the Nash-bargaining solution, except for the extreme cases of $\alpha = 0$ and $\alpha = 1$, where the respective participation constraints have to be added to the problem.

model for different reservation utilities $\bar{u}$, (ii) the alternating offer game for different discount factors δ_p, δ_a, and (iii) the Nash bargaining game for different bargaining power coefficients α. However, while changing α always changes the equilibrium contract, this is not the case for changes in $\bar{u}$ and in δ_p, δ_a.

5 Conclusion

In this note we have analyzed three approaches to account for bargaining power in a moral hazard framework, each of them leading to the same set of contracts. Nevertheless, their usefulness will vary depending on the particular problem under consideration. For example, solving the alternating offer game may be quite cumbersome. Similarly, measuring changes in bargaining power by $\bar{u}$ is unsatisfactory if one wants to understand the impact of bargaining power on equilibrium utility.

Moreover, for sufficiently low levels of the agent's bargaining power, increasing it marginally does affect the equilibrium in the Nash bargaining game, but not in the P-A model and in the alternating offer game. This is in clear contrast to standard bargaining problems without moral hazard. Formally, it arises from the fact that the agent's threat point in the Nash bargaining game and his impasse point in the alternating offer game do not belong to the domain of the function $\pi(u)$ that describes the constrained Pareto frontier $\mathcal{P}$ (see, e.g., Muthoo 1999, 60).

There is a wide range of potential applications. For example, our moral hazard model could represent a firm-worker relationship, where the worker's effort is non-contractible. In this case the appropriate approach to account for bargaining power seems to depend on the particular framing of the problem. If we investigate the implications of a decreasing power of labor unions, then the most natural modelling approach would probably be the Nash bargaining solution or the alternative-offer game. By contrast, changes in the social security system would most naturally be modelled as changes in the worker's reservation utility. Nevertheless, both will have similar effects in our model, which suggests that reducing the worker's bargaining power too much may reduce overall efficiency.

References

Balkenborg, D. (2001). How liable should a lender be? The case of judgement-proof firms and environmental risk: Comment. *American Economic Review 91*(3), 731–738.

Binmore, K., A. Rubinstein, and A. Wolinsky (1986). The Nash bargaining solution in economic modelling. *RAND Journal of Economics 17*(2), 176–188.

Demougin, D. and C. Fluet (2001). Monitoring versus incentives. *European Economic Review 45*, 1741–1764.

Inderst, R. (2002). Contract design and bargaining power. *Economics Letters 74*, 171–176.

Kim, S. (1997). Limited liability and bonus contracts. *Journal of Economics and Management Strategy 6*, 899–913.

Milgrom, P. (1981). Good news and bad news: Representation theorems and applications. *Bell Journal of Economics 12*, 380–391.

Mookherjee, D. and D. Ray (2002). Contractual structure and wealth accumulation. *American Economic Review 92*(4), 818–849.

Muthoo, A. (1999). *Bargaining Theory with Applications*. Cambridge: Cambridge University Press.

Pitchford, R. (1998). Moral hazard and limited liability: The real effects of contract bargaining. *Economics Letters 61*(2), 251–259.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica 50*, 97–109.